A Work Project, presented as part of the requirements for the Award of a Master Degree in Economics from the NOVA − School of Business and Economics.

Out in the dark

Measuring the gay and lesbian wage gap

Patrícia Sofia Pinto e Filipe, 3079

A Project carried out on the Master in Economics Program, under the supervision of:

Professor Pedro Portugal

January 3, 2017

Abstract

Data insufficiency hampered the academic research of discrimination based on sexual orientation. This is a particular concern in Portugal, a country that in spite of the strong legal recognition of homosexuals still scores low in their acceptance. Resorting to a self-designed web-survey, this study provides for the first time academic investigation of homosexuals in Portugal and contributes with an evaluation of wage discrimination in the primary employment. The empirical results point in the direction of absence of discrimination, but there is imprecise and small evidence that some homosexual individuals may be subject to discrimination.

Keywords: sexual orientation, wage, discrimination, Portuguese labour market.

Acknowledgements: *I thank everyone who made the study of this (very challenging) topic possible. To my supervisor, Professor Pedro Portugal, for his endless guidance and patience. To the Professors and Economists who offered me some of their time and advice. To my beloved family, for the continuous stimulus, motivation and support, especially my aunt, Estela, for her example, enthusiasm and assistance. To Inês for her help and care. To ILGA Portugal, and Marta Ramos, that provided assistance in the assembly of the web-survey, and to everyone who helped and supported me throughout this journey. Finally, I express my utmost gratitude to the organisations and individuals who sided with me in this cause, answering and sharing my web-survey. You allowed the study of this sensitive and concealed topic to be pursued, and without you, this pioneering work would be condemned to fail.*

Contents

List of Figures

List of Tables

1 Introduction

Labour market discrimination for lesbians and gays entered the spotlight of academic research only about 22 years ago with the study of Badgett (1995). Although there is a growing interest in such topic, it is often unstudied because of insufficient data (Klawitter and Flatt 1998). In the US, there are multiple surveys that enable research on this topic (General Social Survey, Third National Health and Nutrition Examination Survey, and National Longitudinal Study of Adolescent Health). In Europe, only some countries have available data on sexual orientation. Examples are: the Netherlands, with a database with information on sexual orientation since 1996 (Plug and Berkhout 2002); and Sweden, that legalised homosexual partnerships through civil unions earlier in 1995, therefore granting enough available data to study this minority (Ahmed and Hammarstedt 2010).

As of now, the existing literature repeatedly finds labour market gaps between homosexuals and heterosexuals, regardless of the country. Although research varies in the identification strategy for same-sex partnering individuals, the gaps are consistent: males are penalized in employment and earnings, while lesbians observe ambiguous effects. However, the magnitude of gaps varies widely and with diverse explanations: discrimination, occupational segregation, gender roles or other individual characteristics.[1]

On Lesbian, Gay, Bisexual, Transgender and Intersex (LGBTI) academic research, Portugal remained a concealed country, until now, due to the absence of available data. In the labour market, Portugal implemented the Employment Equality Directive (2000/78/EC) and is ranked by Rainbow Europe (2017) as the 6^{th} best in the world regarding respect for human rights and full equality. In spite of this signal that equality and non-discrimination in employment may be already achieved (ILGA Europe 2017), Valfort (2017) reports that, in Portugal, the acceptance of homosexuality is below the OECD average. Furthermore, the web-survey dataset suggests that many homosexuals working in Portugal are either still closeted in the current workplace, in fear of disclosing their sexual orientation, or suffer from discrimination.

To understand whether this fear of discrimination is justified, I undertake in this study the first ever academic research on homosexuality in Portugal. Section 2 highlights the theoretical expla-

[1] Further developed in Section 3.

nations of labour market discrimination based on sexual orientation and Section 3 the existing literature on the topic. Although it is a research field with many actions to be taken, I introduce the topic by estimating differences in wages between homosexuals and heterosexuals. To overcome the data obstacles that have discouraged previous research, this research uses both a national survey and a web-survey that was projected specifically for this work, described in Section 4. Section 5 provides the identification strategy adopted, followed by Section 6 on the data treatment and econometrics strategy used, drawn from the literature. Results are discussed next, in Section 7. Lastly, Section 8 exposes the main conclusions, the first to be drawn in the Portuguese labour market for individuals with a homosexual orientation.

2 Theoretical Explanations on Homosexuality Discrimination

Several theoretical explanations were adopted ad hoc to explain labour market gaps concerning homosexual and heterosexual individuals. I will address the ones referred in the literature: statistical and taste for discrimination, and household specialisation.

Taste for discrimination (Becker 1957) is when employers dislike homosexuals and, considering they maximise utility instead of profits, they either offer worse employment conditions to homosexuals or prefer heterosexual employees even if with lower productivity or requiring higher wages.[2] In the presence of market failures, this situation can be sustainable in the long term and induce segregation of homosexuals into certain occupations and industries. Furthermore, workers and costumers that discriminate homosexuals may also provoke segregation. Plug et al. (2014) suggest prejudiced straight and homosexual workers choose different occupations, thereupon proposing prejudice-based segregation.

Statistical discrimination, based on incomplete information models (Phelps 1972; Arrow 1973), entail that employers infer the productivity of each homosexual employee through the average productivity stereotype associated with gays and lesbians. This general inference arises because employers' have incomplete information respecting each individual employee, and hence base their decision on a belief of the average productivity of the group. For homosexual individuals, statistical discrimination is related with the heteronormative gender roles assumed in

[2] Additionally, Becker (1976) also makes the point that, because homosexuals are a smaller proportion of the population, discrimination imposes costs on homosexuals rather than benefiting heterosexuals.

society: it is prejudicial for gays because of less masculine stereotypes, but beneficial for lesbians as more productive and dedicated to the labour market. Even with inaccurate stereotypes, a statistical discriminant behaviour may become self-fulfilling and drive the gay employee to underinvest in human capital or accept fewer returns for the investment. A field experiment in which participants take the role of employers (Baert 2017), showed that more risk-averse employers discriminate homosexual males more.

Another explanation other than discrimination for homosexual labour market gaps pursue Beckers (1981) Treatise on the Family: same-sex households do not have the comparative advantages of different-sex households and hence will not specialise in labour market nor domestic tasks. The absence of the abovementioned comparative advantage may change investments in labour market human capital and, consequently, affect wages. For example, gays are assumed to engage comparatively more in the household human capital, whilst lesbians have higher incentives to invest in labour market human capital. In the US, Jepsen and Jepsen (2015) found similarities between gay and straight couples in earnings differentials, but lesbians shared household tasks more equally and evidenced smaller within couple differentials in earnings and hours worked.

3 Literature Review

Bagdett (1995), acknowledged by the literature as the first econometric study, used the US General Social Survey (GSS) and estimated a gay and bisexual gap of minus 11% to 27% in earnings for equally productive individuals compared to heterosexual. Lesbians showed no statistically significant differences. Later research supported this gay negative gap and added a lesbian premium with GSS. Explanations ranged from income effects within the household (Berg and Lien 2002); gay discrimination and higher incentives to invest in labour capital by lesbians (Black et al. 2003); and segregation into occupations for open homosexuals or bisexuals (Blandford 2003). Using US Census, Klawitter and Flatt (1998) showed that the negative gap is smaller amongst high earners, the lesbian premium can be due to superior human capital accumulation, and the gay penalty remains unexplained; Antecol et al. (2008) estimated a gay penalty in comparison with married men but a gay premium to those only cohabitating with a female partner, and a premium for lesbians compared to married and male partner cohabitating

women; and Allegretto and Arthur (2001) found a negative earnings gap for gays mostly (but not fully) explained by marriage. Carpenter (2007), using the Third National Health and Nutrition Examination Survey, corroborated the previously found gay penalty. Sabia (2014), controlled individual heterogeneity using the National Longitudinal Study of Adolescent Health on young adults and still showed a penalty for gay and bisexual men. For women, it was not statistically significant. The gay penalty is hence persistent in all US databases, and the lesbian premium is ambiguous and depends on the data and specification used.

In Sweden, Ahmed and Hammarstedt (2010) used the Longitudinal Integration Database for Health Insurance and found a gay 10% to 15% disadvantage in earnings, more accentuated in non-metropolitan areas. The gap for lesbians was not significant. Ahmed et al. (2013) used Census data and suggested there is gay discrimination and segregation into certain occupations, particularly amongst high earners. Lesbians were estimated to have an advantage at the top of the earnings distribution, possibly due to non-traditional gender roles. Moreover, the advantage was lower in the private sector, a likely result of discrimination.

Using the Household Income and Labour Dynamics in Australia, Sabia et al. (2017) estimated a penalty for gays, partially explained by differential earnings growth, and a lesbian premium caused by greater work intensity. Openly gays had greater penalties, suggesting discrimination. La Nauze (2015) also found a premium for lesbians (0 to 13%) and a penalty for gays (8 to 18%), robust to personality traits and hence suggestive of discrimination. Buser et al. (2015) used a Dutch online survey panel and observed that gays compete less than other men, but lesbians compete as much as other women. The lower gay competitiveness either explains a proportion of the gay penalty or is endogenous to earnings since better-paid positions are assumed to have higher competitive levels but gays can be discriminated in career promotions.

Nevertheless, there are also contradicting findings: Carpenter (2008), using the Australian Longitudinal Survey of Womens Health estimated a lesbian penalty for young lesbians in Australia, paired with higher stress levels and work dissatisfaction; Carpenter (2005), using the GSS for California, found no earnings effect for homosexual workers, but a 10% negative difference for bisexual individuals relative to heterosexuals; and Plug and Berkhout (2002) used a yearly survey of Dutch graduates and concluded no discrimination for young and highly educated

graduates entering the labour market.

These contradicting results not justified by country specificities show the weaknesses of these analyses, as there are persistent problems with sample selection and unobserved heterogeneity.

4 Data

Two different databases were used in this research: Inquérito ao Emprego, the only Portuguese database with detailed employment information that allowed the identification of partnered homosexuals, and a web-survey created specifically for the purpose of this research.[3]

Inquérito ao Emprego is a national survey executed by Instituto Nacional de Estatística (Statistics Portugal) that focuses on collecting labour market information of households in Portugal. The data collection has a quarter periodicity and happens at the household level, where information is gathered individually for the ones living there permanently. Individuals are followed during six quarters, providing a maximum of six observations with three months interval per individual, in panel data format. Whenever a household leaves in between the sampling full length, it is replaced by another from the same area. The selection of households is random and representative of the population in Portugal.[4] Individuals employed by Statistics Portugal are responsible for conducting the interviews during the data collection, registering answers with a Computer Assisted Personal Interviewing in all periods unless the individual agrees on a Computer Assisted Telephone Interviewing after the first interview. Such method guarantees the reliability of the data. From the 1^{st} quarter of 2011 (Q1 2011) up to the 2^{nd} of 2017 (Q2 2017), the period covered in the research, there are 1,056,584 observations available. Notwithstanding, only 154 homosexuals are identified.[5]

In the absence of enough available data to investigate the LGBT minority in Portuguese databases, I resorted to another method: an online data collection based on self-reporting through a web-survey. When designing the web-survey, I focused on questions identical to the ones available in Inquérito ao Emprego and relevant to the study.[6] Such method was intended to minimise disparities between answers, enable the merger of the two datasets and narrow comparability

[3] The first sample of homosexual individuals is similar to the Allegretto and Arthur (2001) but, to my best knowledge, the second sample is not used in the more influential literature.

[4] It follows stratification through NUTS III, and is in accordance with article 3 of the 577/98 European Union regulation and national directives.

[5] Correspondent to 51 individuals that are repeatedly observed.

problems.

Together with the questions from Inquérito, questions specific to the LGBT community were added to tackle literature limitations and grant a more in-depth analysis. Accordingly, questions were created based on the information provided in personal interviews with LGBT experts and granted four supplementary advantages to the analysis: (1) clear distinction between individuals who disclose their sexual orientation at work; (2) understanding the reasons for those that remain closeted; (3) perceiving the self-disclosure behaviour in the workplace compared to other co-workers disclosure decisions; and (4) evaluating commonly reported discriminatory situations in the workplace. Later, these additional questions were approved by LGBT experts and academics experienced in designing surveys, assuring they were sensible to the vulnerabilities of the LGBT minority, but at the same time academically reliable and capable to be analysed in an unbiased manner. The final version of the web-survey is available in the Appendix.

The diffusion started in October 2017. The main network was social media channels, using both personal accounts, with a total of 73 shares, and organisational accounts of 10 LGBTI Portuguese institutions, including the islands.[7] The final dataset had 1211 observations, 782 without inconsistent answers.

Against all my efforts to share the web-survey as far and broadly as possible, the shortcomings of the diffusion method may still limit the population covered. First, although I shared the web-survey with diverse LGBTI organisation, this limitation can provide biases of more affirmative individuals following LGBTI organisations. Second, despite my strong efforts to induce as many contacts of my personal network to share the survey, it can still be biased towards individuals engaging in social media and connected through the contacts of my personal network.

5 Identification Strategy

To overcome the lack of data problem, researchers use different mechanisms to trace homosexuals (Ahmed and Hammarstedt 2010). In this study, the focus is on two different methods that

[6] General labour market questions are a replica from the ones available in Inquérito ao Emprego metafiles, for example work region, wage, industry of work, occupation, full-time status, job search, tenure and the remaining used as controls in the regressions.

[7] I contacted all organisations that were mentioned in the Portuguese section of the EU LGBT Survey (2012) and others that, to my knowledge, are established in Portugal and have a wide coverage.

lead to distinct samples: individuals recorded in Inquérito ao Emprego that report cohabiting with a same sex partner (Carpenter 2004); and self-reporting from the web-survey.

In Inquérito ao Emprego, 35 homosexual males (0.09% of males) and 16 homosexual females (0.04% of females) that qualified to the definition employed were identified. The percentage identified is significantly lower than the 5.9% estimate for Europe's self-reported homosexual population provided by Dalia research (2016).[8] Considering Inquérito ao Emprego data collection process restraints anonymity (the first interview is always personal), it incentivises homosexual individuals that want to protect their identity to hide their sexual orientation, not identifying the other same-sex household member as partner.

Inquérito ao Emprego, therefore, suffers from two problems when identifying the sample of homosexual individuals: underreporting of homosexual individuals cohabiting with a same sex-partner that choose to not identify the partner as such; and measurement error because the Inquérito ao Emprego construction prevents the identification of single homosexuals or homosexuals not cohabiting with the partner. These issues are a problem for the identification of homosexual individuals, and simultaneously aggravate the measurement error in the counterfactual (straight individuals) used in our analysis.

Concerning the web-survey, homosexuals were identified through the self-reporting question *How would you identify your sexual orientation?*. The final sample of individuals reporting homosexuality and fulfilling the robustness checks comprises 155 homosexual males (67.8% of males) and 61 homosexual females (25.0% of females). Since the collection mechanism is fully anonymous and individuals have the choice to answer the web-survey and to decide which questions to answer (no question was mandatory to proceed), it does not provide incentives to misreport. I, therefore, assume all individuals reported truthfully their sexual orientation. This advantage explains why individual interviews were avoided, as conducted in Inquérito ao Emprego, weighting the trade-off between underreporting and fewer observations.

Regarding the counterfactual group available in the web-survey, it was created by filtering indi-

[8] Dalias estimate should be closer to the true proportion of homosexuals in the population because it is a census-representative sample and is collected through surveys via web, preserving anonymity. Although the self-reporting estimation can be far away from the true proportion of homosexuals in the population, I assume individuals more willing to self-report their sexual orientation are also the ones more likely to disclose their sexual orientation at work (Badgett 1995). Since homosexuality is unobservable, the individuals that are more relevant to the analysis are the ones openly homosexual at work.

viduals that selected heterosexual in the sexual orientation self-reporting question.[9]

Be as it may, it must be emphasised that the results estimated using the web-survey only represent the population reached, which may enlarge estimation biases through self-selection. Concomitantly with the diffusion mechanism bias, it is biased because it is not random nor representative of the population working in Portugal, unlike Inquérito ao Emprego. Conjointly, there is possible self-selection in opening the link and fully answering the web-survey caused by unobservable characteristics that may influence earnings as well.

Inquérito ao Emprego, by virtue of homosexuality unobservability, is biased towards individuals that are more likely to voluntarily report their homosexuality. In turn, such willingness to disclosure can be related with unobservable characteristics that may also impact earnings. Moreover, because only homosexual couples are identified, there may exist other unobservable characteristics related to being in a relationship that also affect earnings. Such unobservable characteristics may cause omitted variable bias.

6 Empirical Strategy

The Inquérito ao Emprego dataset was already clean when made available by the Bank of Portugal. Nevertheless, I had to clean the web-survey dataset and then merger the two. Since the respondents of the web-survey did not have assistance when answering it, I conducted tests for duplicate and falsified answers, maintaining the measurement error at a minimum. The first mechanism was cross-checking the time the respondent used to complete the survey and the starting and end time of each observation.[10] Afterwards, a consistency evaluation between all questions responded for extreme answers of hours worked and wages was conducted, particularly regarding occupation, industry, and the situation as self-employed, employer or employee. Moreover, additional controls were used to ensure that there were no observations with senseless reporting. The aforesaid controls were: consistency between sexual orientation and relationship status; age consistent with education, experience and tenure; living region in accordance with professional situation and work region; and, when available, match the job description with

[9] In any case, a robustness check for sexual orientation was used in individuals that reported being in a relationship.

work industry, occupation and wage.[11]

After ensuring a clean web-survey dataset, an exact matching on covariates approach, the most reliable matching method according to Cameron and Trivedi (2005), is used. This method relies on creating a counterfactual group, with observations exactly equal to the ones in the treatment group in all observables but treatment, for a good-sized number of treated observations (common support condition).[12]

In light of the sampling plan limitations described in the previous section, particularly the non-randomness of the web-survey sample, this approach provides the advantage of estimating more precise gaps in wages of homosexuals and heterosexuals.[13]

As described in Blackwell et al. (2009), the matching procedure used assigns all observations to a certain stratum, representative of a unique combination of the covariates chosen: education, age group and full-time employment. Subsequently, treated and untreated observations are matched if assigned to the same stratum.

A preliminary result for wage differences caused by sexual orientation is assessed with the average treatment effect on the treated (ATT). Since homosexuality, the treatment, is assumed to be exogenous to human capital accumulation and labour market decisions, I obey the Conditional Independence Assumption.[14] However, because the exact matching is very demanding on the number of observations available, the covariates chosen were limited only to the main determinants of earnings as suggested by earnings-age profile (Mincer 1974). In my view, such analysis hence provides an upper bound of discrimination in the presence of segregation, but this is not the only estimation procedure because of the few imposed covariates to match the observations.

The ATT is here computed as a simple average of the outcome of interest due to treatment. Since there are different numbers of individuals per stratum, and not all strata contain homosexual and heterosexual observations, the relevant strata were selected to compute the within stratum wage

[10] For individuals that initiated a new questionnaire at the same time another was finished, I carefully evaluated whether answers were consistent throughout the questionnaire.

[11] The correction of work industry and occupation through the job description was handled by an independent person without economic knowledge, in accordance with Statistics Portugal guidelines, to grant unbiased estimations and avoid concerns over induced spurious results.

[12] By depending on observable characteristics only, it assumes that unobservables do not influence the variable of interest.

[13] Although the control group (heterosexuals) is subject to the measurement error in sexual orientation, it presents itself as the best counterfactual to the web-survey homosexuals because it is representative of the population working in Portugal.

[14] This assumption assumes that sexual orientation is independent of unobservable characteristics: $y_0, y_1 \perp D \mid x$, where $D = 1$ if homosexual and $D = 0$ if heterosexual.

difference between homosexual individuals and correspondent heterosexuals: $TE_i = y_{i1} - y_{i0}$. Later, an average of all the differences is used to estimate the sample average treatment on the treated: $SATT = \frac{1}{n_T} \sum_{i \in T} TE_i$.

The common support condition allows not only the primary analysis of wage differences but enables also the merge of the two sampling plans, heterosexual observations form Inquérito ao Emprego and homosexual observations from the web-survey, to be used in following analyses.

Subsequent analyses apply a OLS regression strategy for cross-sectional data, as used in most of the literature reviewed. The base model is an augmented version of the traditional earnings equation derived by Mincer (1974).

Mincer model was first derived to capture schooling rate of return on earnings (Mincer 1958), where individuals decided the education level that maximised the present value of their life-time earnings. Nevertheless, because schooling is only a primary investment in human capital, Mincer added the impact on earnings of investments in human capital throughout working life. Suggesting that the proportion of earnings given to investment in human capital decreases linearly with experience, the final specification is given by: $lnY_t = a + b_1 s + b_2 x_t + b_3 x_t^2 + w$, where lnY_t is the log of earnings, s is schooling, x_t is the accumulated investment in human capital of workers, given by potential experience, and w the error term.[15][16]

Since in this study the research question is not directly related to investment in human capital returns, Mincer's equation is a starting point to remove the impact of other variables that explain earnings, isolating the impact of homosexuality. Therefore, the variables proposed by Mincer are used as *potential experience* $= age -$ *years to complete the school level* $- 6$ as in Allegretto and Arthur (2001); and education through level corresponding dummies, as recorded in Inquérito ao Emprego and the web-survey. Potential experience, albeit limited because it disregards interruptions in employment and investment in human capital, is more representative of the trade-off between schooling and work experience.[17] Education through dummies concomitantly controls for possible *sheepskin effects* and allows the estimation of different experience-earnings profiles that are not parallel to individuals with different education, as in

[15] Previous to Mincer, Becker (1964 and 1967) proposed the theory of optimal allocation of investment in human capital over the lifecycle, in which individuals invest more only at younger ages because of the eagerness to shift from learning towards earning activities. Also, because individuals only invest in human capital while the rate of return of the investment is higher than the discount rate, the investment increases up until the return is equal to the discount rate and falls afterwards.

[16] *potential experience* $= age -$ *estimated age of school completion*

Mincer (1974).[18]

Furthermore, vectors of additional controls to Mincers specification are included to isolate their effects on the wage and reach a better measure of the impact of sexual orientation:

$$lnY_i = \alpha + \delta LG_i + \beta_1 \mathbf{X}_i + u_i \quad (1)$$

For the specification given by Equation 1, three separate datasets are used: a combination of Inquérito ao Emprego heterosexual and web-survey homosexual observations; observations from Inquérito ao Emprego only; and observations from the web-survey only. LG_i is the parameter of interest and δ measures the impact of being homosexual, on average, *ceteris paribus*. In all estimations, the dependent variable, lnY_i, is the net wage of the primary job, since it is the only wage measure available in Inquérito ao Emprego. In case of differential treatment between homosexual and straight individuals in similar jobs, this specification identifies the differences in offered base wages only, disregarding career progressions and bonuses. The use of logarithms on earnings is consistent with both the units of investment being in time and not monetary units (Mincer 1975), and the accurate fit of log-linearity for the major wage distribution of Mincers sample. Additionally, it simplifies the interpretation of small coefficients.

$\mathbf{X}_i$ is a vector of control variables and β_1 is a vector of regression coefficients containing a coefficient for each of the controls used. The remaining unobservable effects not captured in the controls are in the error term, u_i, which is assumed to have zero mean and to be uncorrelated with the explanatory variables.[19]

These abovementioned controls include socio-demographic characteristics such as nationality and marital status. Marital status bias for homosexuals may result if they have a lower probability of getting married. Other controls included are human capital and work specific. Education and potential experience were justified by Mincer's equation but can also be correlated with sexual orientation if, for example, homosexuals have more difficulties finding a job (Weichselbaumer 2003 and 2015; Drydakis 2009; Tilcsik 2011). Employment region is also included because of results that estimate homosexuals self-selection into certain areas (Black et al. 2007).

[17] In the absence of better data, the years spent to attain the education level are considered as the minimum years required to the degree obtained.

[18] The experience-earnings profile parallel to all individuals is an assumption of the original Mincer model but that it is now being challenged by recent literature.

[19] If this unobservable effects are correlated with the coefficient of interest, for example through sample selection, δ, it may cause an omitted variable bias

Workplace industry (21 categories) and occupation (10 categories) controls aim to remove potential effects caused by segregation of homosexuals into certain industries and occupations (for example Badgett 1995). Since industry and occupation can be endogenous to sexual orientation (La Nauze 2015), including this controls provides only a lower bound estimate of discrimination when there is negative segregation. Hours worked and full-time are also used in the literature and thus included in our models (for example Plug and Berkhout 2002).

In addition, controls for whether the individual is looking for another job, the size of the work place, type of work contract, professional situation (military, civil employee or student) and tenure are included as well. Although these variables are not common in the literature reviewed, they might be correlated with wages and sexual orientation, and to estimate the impact of wage discrimination their effects on wages should be removed. Although some of the aforementioned variables can also represent forms of discrimination, the scope of this research is to isolate wage discrimination only.

Finally, specific to Inquérito ao Emprego Q1 2011 to Q2 2017 sample, quarterly fixed effects to control for time trends and the business cycle are also used. In the other samples I only consider observations corresponding to Q2 2017 of Inquérito ao Emprego and the recently collected web-survey, hence do not require such controls.

The abovementioned specification allows testing for statistical discrimination and taste for discrimination through wage differentials. Notwithstanding, if sexual orientation is not observable to the employee, there is an additional measurement error that can compromise the unbiasedness of the estimators. If, for example, a homophobic employer perceiving a homosexual individual as heterosexual may refrain from discriminating. To control for this bias, often a limitation in the literature, another specification is followed, Equation 2, that uses disclosure in the workplace dummies instead of sexual orientation:

$$lnY_i = \alpha + \gamma_1 FD_i + \gamma_2 PD_i + \gamma_3 ND_i + \beta_1 \mathbf{X}_i + v_i \tag{2}$$

Full-disclosure in the work place to all co-workers and superior management is given by FD_i and captured in γ_1. PD_i is partial disclosure to some co-workers, estimated in γ_2, and ND_i is no disclosure at all given by the coefficient γ_3. In this specification, only the models that include web-survey homosexual observations are used, since Inquérito ao Emprego does not provide a

break-down of disclosure status per homosexual observation.

In all specifications, a linear relationship between log wages and covariates, consistent with the OLS structure, is considered.[20] Although the web-survey homosexual sample is not random, matching grants that there is a good common support at least in Inquérito ao Emprego to proceed with the analysis.

Conditional on the variables included in the regression, the selection of unobservable characteristics is considered independent of sexual orientation. Moreover, because Inquérito ao Emprego design prevents a full identification of homosexuals willing to disclose their sexual orientation if single or not cohabiting with their partner, there is a potential bias that cannot be solved, and it is a limitation considered when interpreting the results.[21]

Besides, as Black et al. (2003 and 2007) refers, even though the sexual orientation is exogenous to the individual, it can influence her individual constraints and, consequently, choices. Hereinafter, to complement the former analysis, a Blinder-Oaxaca decomposition to estimate if there are differences in endowments or regression coefficients of individuals with distinct sexual orientations is included. As described by Jann (2008) and used in La Nauze (2015) and Antecol et al. (2008), the simple decomposition is given by:[22]

$$\overline{lnY^{LG}} - \overline{lnY^{ST}} = \beta^{LG}(\overline{X^{LG}} - \overline{X^{ST}}) + (\beta^{LG} - \beta^{ST})\overline{X^{LG}} + (\beta^{LG} - \beta^{ST})(\overline{X^{ST}} - \overline{X^{LG}}) \quad (3)$$

where LG signals homosexual individuals and ST straight ones. $\overline{X}$ represents the average of the controls used.

This method is based on two distinct linear regressions of Equation 1, one for homosexual individuals only and the other for heterosexual individuals. After the individual regressions, Equation 3 is formulated into a threefold decomposition derived from the first estimations: (1) the difference in the homosexual and heterosexual average endowments; (2) the change in the homosexual mean outcome due to differences in the coefficients; and (3) the concurrent difference in endowments and coefficients. The term $(\beta^{LG} - \beta^{ST})\overline{X^{LG}}$ captures the discrimination because it captures the difference in regression estimated coefficients to the same characteristics,

[20] $E[y \mid x] = \mathbf{x}'\beta$, such that $y = E[y \mid x] + u$ and $y_i = \mathbf{x}_i'\beta + u_i$. Furthermore, standard errors are corrected using Huber, White and sandwich estimators and standard error corrections also use clustered sandwich estimators to accommodate the within correlation of the error terms.

[21] Despite the alternative econometric adjustments for the estimators, the nature of the samples and the number of observations in the treatment group prevents using further econometric corrections.

[22] The equation used in adapted from Jann (2008).

for individuals with different sexual orientations. However, because there are few observations in homosexual samples and the web-survey heterosexual sample, using individual regressions for each sample of homosexuals and heterosexuals separately may hamper the number of degrees of freedom in each regression. To account for this problem, the controls used in the Blinder-Oaxaca decomposition are the minimum necessary: experience, education levels and full-time status. These are the same covariates used in the exact matching.

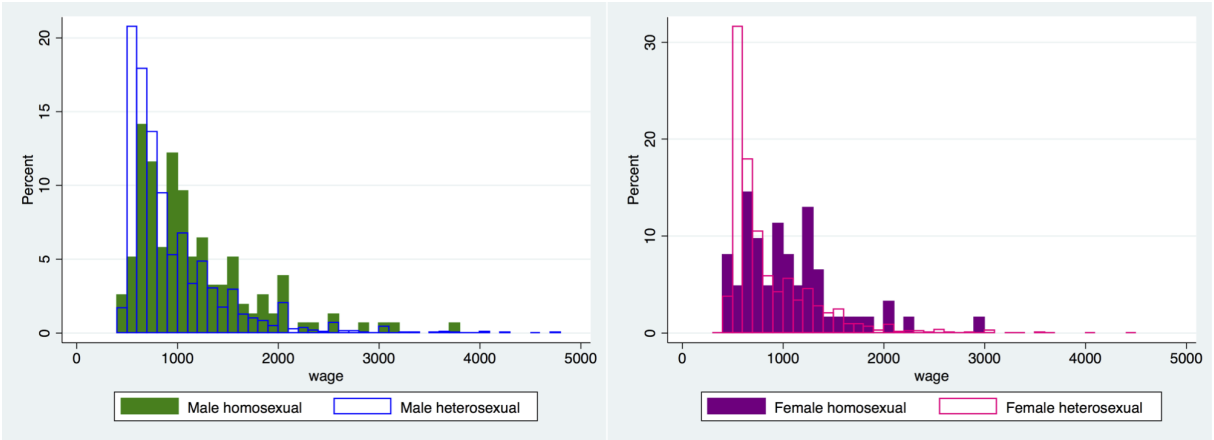
7 Empirical Results

7.1 Descriptive Statistics

With a non-random and non-representative treated group, outliers were excluded from both the treatment and the control group. The net wage upper bound was set at 5000€, up to where there is the main concentration of observations and the lower bound was computed as *year of the interview minimum wage*$*0.8 * 0.89$.[23]

Figure 1 shows the wage distribution of homosexuals from the web-survey and heterosexuals from Inquérito ao Emprego. From Figure 1 it is clear that the web-survey has a higher proportion of homosexual individuals with higher wages than the heterosexuals in Inquérito ao Emprego. This is common to both gay/male and lesbian/female distributions.[24]

Figure 1: Histogram of homosexual and heterosexual wages distribution without outliers.



Source: Homosexual observations form the web-survey and heterosexual observations from Inquérito ao Emprego Q2 2017.

Table 1 displays the mean of the main variables of interest, subdivided by sexual orientation.

[23] The formula accounts for social security taxes, 11%, and the minimum wage of apprentices and interns which is 80% of the minimum wage.

[24] The Figures 3 and 4, displayed in Appendix, benchmark the distribution of homosexuals from the web-survey against the distribution of heterosexuals from the Inquérito ao Emprego in age and education. Figure 2 exhibiting the wage distribution with outliers is also included.

Table 1: Descriptive Statistics of Wage, Education and Age

	Males				Females			
	Inquérito ao Emprego (IE)		Web-survey (WS)		Inquérito ao Emprego (IE)		Web-survey (WS)	
	Straights	Gays	Straights	Gays	Straights	Lesbians	Straights	Lesbians
Wage (euros)	878.75	1109.51	1279.77	1115.46	787.70	905.61	1099.81	1028.01
	(484.78)	(759.18)	(719.56)	(564.51)	(418.74)	(521.25)	(472.64)	(486.44)
Years of Education	9.4	12.4	15.4	15.6	10.6	12.6	15.4	15.2
Age	42	37	34	34	43	40	38	36
Individual/Quarter Observations (N)	136584	101	74	155	147878	49	186	61
Individuals (n)	40214	35	74	155	41776	16	186	61

Source: Means are computed per observation but several observations may correspond to the same individual at different points in time.

As shown in Table 1, individuals from the web-survey are, on average, younger, more educated, and earn higher wages than the ones from Inquérito ao Emprego. Considering Inquérito ao Emprego is representative of the population working in Portugal, the disparities in mean wage, education and age suggest that the web-survey sample is not.

There is also a remarkable gap between homosexual and heterosexual individuals, particularly in Inquérito ao Emprego. In the previously mentioned sample, gays exhibit 26% and lesbians 15% higher wages than their heterosexual counterparts. This result is expected for more educated individuals: gays have, on average, 3 more years of education than straight males, and lesbians 2 more years than straight females. Higher education results were also found by Black et al. (2000) in a US homosexual sample. Furthermore, homosexual individuals from Inquérito ao Emprego are also, on average, 5 to 3 years younger than their counterparts.

In the web-survey, the gaps between individuals of different sexual orientations are less pronounced. Gays display, on average, 13% and lesbians 7% lower wages than their straight counterparts. Differences in education captured in the web-survey are no longer striking, up to 0.2 years, and the sample is also more balanced in terms of age differences for individuals of distinct sexual orientations as homosexuals (on average, only up to 2 years younger). The survey is also disproportionally more representative of Lisbon workers, and of certain activities and occupations. [25]

Concerning homosexuals only, the web-survey captures a higher proportion of homosexuals that assume their homosexuality in the work place compared to those closeted: 41.3% of gays are fully disclosed in the workplace, and 40,0% of lesbians are partially disclosed. With regards

[25] Samples distributions in work region, work industry and occupation are shown in Figure 7, 8, and 9, respectively.

to disclosure in the workplace, there seems to be a correlation between own disclosure and co-workers disclosure, as shown in Table 2.

Table 2: Percentage of disclosed co-workers per disclosure of individuals observed

	Fully disclosed Gay	Partially disclosed Gay	Closeted Gay	Fully disclosed Lesbian	Partially disclosed Lesbian	Closeted Lesbian
Fully disclosed co-workers	69.8%	22.7%	7.5%	53.8%	30.8%	15.4%
Total Observations	41.3%	38.9%	19.8%	30.8%	40%	29.2%

Source: Homosexual observations form the web-survey.

The proportion of male workers with fully disclosed co-workers that disclose their homosexuality is 69.8%, while closeted is just 7.5%. For lesbians, 53.8% of the ones with fully disclosed co-workers also disclose their homosexuality while only 15.4% remains closeted in the workplace. This can be related with a more environmentally friendly workplace, since several homosexuals come out, or because a more environmentally friendly workplace causes more homosexuals to come out. Such proportion in the survey can provide estimates more representative of individuals willing to disclose their sexual orientation in the workplace, than in turn can reflect less discriminatory workplaces.

Nonetheless, from the 61.6% that did not fully disclose their homosexuality at work, 25.9% are afraid of not being accepted, 39.2% are scared of moral harassment like jokes or prosecution, and 23.8% fear a career penalty. Even more striking is that 24.5% of individuals reported suffering from moral harassment in the current workplace. Although I do not have enough observations to comprehensively study this type discrimination and its' impact, it is an alert of psychology theories. Such discrimination may negatively impact job and career attitudes (Ragins and Cornwell 2001), and result in less promotions, compensations, and lower productivity (Badgett 1996) that further impact the wages negatively.

7.2 Matching and Regression Results

7.2.1 Wage gap with Exact Matching

On the regression results, first there is an exact matching on education, age group and full-time status. As displayed in Table 3, the proportion of exact matches in the three covariates chosen is fairly high and the wage differences are negative but not statistically significant.

Moreover, there are, on average, 17 males identified as heterosexual matched to each gay obser-

Table 3: Exact matching in education, age group and full time status

	Gays	Lesbians
Average Treatment on the Treated (ATT) - wages in €	**-1.82**	**-7.05**
	(17.30)	(26.79)
ATT proportion of the median wage	-0.2%	-0.9%
Total Observations Matched (N)	2486	2248
Homosexual Observations Matched	145	61
Percentage of Homosexual Observations matched	93.5%	100%

Source: The data matches homosexual observations from the web-survey and heterosexual observations from Inquérito ao Emprego of Q2 2017. Standard errors are displayed in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

vation, and the proportion of gays matched with an heterosexual counterpart is 93.5%. Regarding females, there are, on average, 37 heterosexual match females to each lesbian and 100% of lesbian observations are matched with at least one heterosexual female. Thereupon, the high proportion of homosexual observations matched and the significant number of correspondent observations matched from the control group ensures the existence of a good enough counterfactual to the homosexual observations. From this strong match follows the conclusion that there is no statistically significant discrimination, or at least a very small and imprecise penalty.

7.2.2 OLS estimate of the homosexual orientation impact on wages

Following is the OLS results, which allow inference with an additional set of controls not used in the matching. To isolate the homosexuality impact on earnings as much as possible, controls that may influence earnings *per se* and that could bias the estimated impact of the LG_i variable, $\widehat{\delta}$, were used, as mentioned in Equation 1.

Regardless of the sampling plan and conditional on the controls used, there is no statistically significant impact in the current job wage of being gay or lesbian. According to Table 4, the limited evidence provided by the estimates, conditional to all controls included, suggest a slight impact of being gay from -2.6% up to a 7.7% premium on the wage, which contradicts literature findings that consistently report significant or limited penalties. The estimated negative impact of being gay may be explained by household effects, consistent with Berg and Lien (2002).[26]

Concerning lesbians, Table 4 presents a small positive effect of homosexuality between 4.4% and 5.7%. For lesbians, both the lack of statistical significance, suggestive of an imprecise effect, and the limited premium are in accordance with the literature reviewed, which provides

[26] In the specification $lnY_i = \alpha + \delta LG_i + \sigma SSC_i + \beta_1 \mathbf{X}_i + \beta_2 C_i + z_i$, accounting for possible household effects on the wages, there are imprecise positive effects of being homosexual after the inclusion of household effects, as exhibited in Table 11 available in the Appendix. Such evidence suggests that the gay negative gap found in column 6 may be driven by the household effect of having a cohabiting gay partner.

Table 4: OLS estimation impact of homosexuality on wages

Variables/Samples	Homosexual (WS) & Straight (IE)		Homosexual & Straigth (IE)		Homosexual & Straight (WS)	
	(1)	(2)	(3)	(4)	(5)	(6)
Males						
Gay	0.189***	**0.045**	0.165*	**0.077**	-0.114*	**-0.026**
	(0.04)	(0.04)	(0.10)	(0.06)	(0.06)	(0.07)
R-squared	0.006	0.602	0.003	0.623	0.014	0.696
Controls		YES		YES		YES
Observations	5797	4940	136685	117325	227	205
Females						
Lesbian	0.224***	**0.047**	0.123	**0.057**	-0.073	**0.044**
	(0.06)	(0.05)	(0.11)	(0.08)	(0.06)	(0.05)
R-squared	0.003	0.724	0.004	0.739	0.006	0.716
Controls		YES		YES		YES
Observations	6220	5595	147898	133744	244	222

Source: Columns 1 and 2 use homosexual observations from the web-survey and heterosexual observations from Inquérito ao Emprego Q2 2017. Columns 3 and 4 use homosexual and heterosexual observations from Inquérito ao Emprego, from Q1 2011 to Q2 2017. Columns 5 and 6 use homosexual and heterosexual observations form the web-survey. The controls used are: nationality, marital status, level of education, experience, work region, full time status, industry, occupation, hours worked, job search, size of the workplace, work contract, professional situation and year-quarter fixed effects. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

mixed evidence between premiums and no effects.[27]

Compared to previous literature, I may be capturing a different bias because the web-survey sample is not representative of the population and involves a selection of individuals distinctive from the ones noted in administrative samples, for example individuals with more vulnerable work situations that restrain from answering the web-survey or that are not reached in the diffusion method, leading to an upward bias in the estimations of column 2. Even though such conclusion is consistent with the web-survey sample of more educated individuals, this option was not considered yet because the more influential literature on wages discrimination does not use self-designed surveys.

Nonetheless, although the estimations in column 6 are consistent with the literature (wage penalties associated with gays and premiums associated with lesbians), the interpretations must be sensitive to the biases in both the homosexual and heterosexual samples. While homosexuals may represent mostly LGBTI institutions, this is not necessarily the case with the heterosexual sample, that could have been mainly assembled with the diffusion from my personal network.

[27] The limited positive estimates from Inquérito ao Emprego decrease when restricting the sample to individuals living with their partner. Notwithstanding, even with the sample restriction, the estimates remain at 4.3% for both gays and lesbians, as shown in Table 10 available in the Appendix. Such result suggests that the positive effect is not attributable to a characteristic common to individuals in a relationship.

7.2.3 Blinder-Oaxaca decomposition of the wage gap

To cement the results estimated previously, a Blinder-Oaxaca decomposition, given by Equation 3, is used. As mentioned in the discussion of the empirical strategy, the term that captures discrimination is the disparities attributable to different coefficients of characteristics.

Table 5: Blinder-Oaxaca Decomposition Results

Variables/Samples	Straight Male vs Gay			Straight Female vs Lesbian		
	(1) Gay (WS) & Straight (IE)	(2) Gay & Straight (IE)	(3) Gay & Straight (WS)	(4) Lesbian (WS) & Straight (IE)	(5) Lesbian & Straight (IE)	(6) Lesbian & Straight (WS)
Total Log Monthly Wage Gap	0.189***	0.165*	-0.114*	0.224***	0.123	-0.073
	(0.04)	(0.10)	(0.06)	(0.06)	(0.11)	(0.06)
Differences in characteristics	0.040	0.125	0.040	-0.141	-0.266**	-0.066
	(0.08)	(0.10)	(0.05)	(0.12)	(0.11)	(0.05)
Differences in coefficients	**-0.019**	**0.054**	**-0.166****	**-0.003**	**0.107**	**-0.031**
	(0.03)	(0.05)	(0.06)	(0.04)	(0.11)	(0.05)
Interaction of both	0.168	-0.014	0.012**	0.368	0.282**	0.024
	(0.07)	(0.05)	(0.04)	(0.12)	(0.11)	(0.04)
Controls	YES	YES	YES	YES	YES	YES
Straight Observations	5642	136584	72	6159	147849	183
Homosexual Observations	155	101	155	61	49	61

Source: The dependent variable is the logarithm of monthly wages. Columns 1 and 4 use homosexual observations from the web-survey and heterosexual observations from Inquérito ao Emprego Q2 2017. Columns 2 and 5 use homosexual and heterosexual observations from Inquérito ao Emprego, from Q1 2011 to Q2 2017. Columns 3 and 6 use homosexual and heterosexual observations form the web-survey. The controls used are: level of education, experience, full time status, and year and quarter fixed effects. Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

In columns 1 and 4 of Table 5, it is showed that homosexuals from the web-survey enjoy statistically significant better average (raw) wages than heterosexuals from Inquérito ao Emprego. Statistically insignificant higher average wages follows for columns 2 and 5, where homosexuals from Inquérito ao Emprego have higher averages than their heterosexual counterparts. However, with homosexuals and heterosexuals from the web-survey (columns 3 and 6), the average wage gap is negative towards homosexuals. In all columns, the wage gap estimates are a direct implication of the wage differences mentioned in Table 1 and Table 4, and hence are consistent with the previously reported results.

Regarding differences in characteristics, the gay sample shows better endowments than straight males in all samples. This positive differences may be explained by a stronger advantage in education despite the disadvantage in potential experience. On the other hand, lesbian samples consistently show worse endowments. This negative gaps on endowments are possibly captured by lower potential experience or a higher proportion of lesbians working part-time compared to

their heterosexual counterparts.

Structural differences, captured by the differences in regression coefficient estimates, suggest that gays from the web-survey experience a negative gap from -16.6% to -1.9% and lesbians from -3.1% to -0.3%. Opposite effects are estimated for the homosexual sample in Inquérito ao Emprego: for gays there is a muted estimate of a 5.4% positive gap and for lesbians of 10.7%.[28] Considering that the decomposition methodology uses separate estimations for homosexuals and heterosexuals, the estimates depend heavily on the individuals represented in each sample. This specificity to the sample considered is particularly relevant for the coefficients on variables that are very sparse, for example certain education levels, or in the case of Inquérito ao Emprego years and quarters. Since there is no guarantee that any of the homosexual samples is representative of the homosexual population, and each sample may be subject to particular biases, external validity may be compromised. Without external vailidity and in light of sample-selection, the estimated opposite effect may reflect differences between the individuals captured in each sample only.

7.2.4 OLS estimate of homosexuality disclosure impact on wages

To validate the previous result and employ the last test on the existence of discrimination, a specification of disclosure in the work place is used. This specification is an improvement to one of the literature most referred limitation: because the web-survey enables the analysis of disclosure effects, it may provide truthful discrimination estimates that are undistorted with the employer perception (see, for example, Plug and Berkhout 2002).

As exhibited in Table 6, there are again no statistically significant effects of disclosure. However, it is noticeable that, for both gay specifications, the impact of being gay is more positive (or less negative) if the individual did not disclose his sexual orientation at work: 6.4% for the merged sample and -1.7% for the web-survey only. This result may be suggestive of statistical discrimination or taste for discrimination, although not plain because it is statistically insignificant and the differences between the imprecise estimations of disclosure and closet in column 2 are very narrow.

For lesbians, there is restricted evidence of lower estimated impacts in partial disclosures, be-

[28] Considering the Inquérito ao Emprego counterfactual of only heterosexual individuals living with their partner, the gay gap decreases to 1.8% but the lesbian gaps remains steady at around 10.4%, as displayed in Table 12 of the Appendix

Table 6: OLS estimation impact of homosexuality disclosure on wages

Variables/Samples	(1) Gay (WS) & Straight Male (IE)	(2) Gay & Straight Male (WS)	(3) Lesbian (WS) & Straight Female (IE)	(4) Lesbian & Straight Female (WS)
Full disclosure	**0.052**	**-0.033**	**0.092**	**0.024**
	(0.05)	(0.07)	(0.08)	(0.08)
Partial disclosure	**0.040**	**-0.023**	**-0.016**	**0.003**
	(0.05)	(0.07)	(0.08)	(0.08)
Closeted at work	**0.064**	**-0.017**	**0.082**	**0.133**
	(0.05)	(0.08)	(0.06)	(0.09)
R-squared	0.602	0.708	0.724	0.7189
Controls	YES	YES	YES	YES
Observations	4939	203	5595	221

Source: Columns 1 and 2 use homosexual observations from the web-survey and heterosexual observations from Inquérito ao Emprego Q2 2017. Columns 3 and 4 homosexual and heterosexual observations form the web-survey. Each model uses either only male observations or female observations. The controls used are: nationality, marital status, level of education, experience, work region, full time status, industry, occupation, hours worked, job search, size of the workplace, work contract and professional situation. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

ing out to co-workers but not managers, compared to no disclosure. The estimated coefficients of column 3 can hence have two different interpretations in view of discrimination theories, assuming discrimination exists. First, if they do not signal their homosexuality, they may not be subject to statistical discrimination that often considers lesbians put higher effort in their job than heterosexual women. In this situation, the impact of being a fully disclosed lesbian is higher. On the other hand, if they do not signal their homosexuality, they may not suffer from taste for discrimination and therefore will not be prejudiced for being lesbian. The latest would justify the higher limited earnings estimated for closeted lesbians also in column 4. Notwithstanding, the hypotheses abovementioned are too strong considering the disparate results between samples. In column 4 there is at least a 13.3% muted gain in lesbians wage, but only for those that choose to remain silent in the workplace (partial disclosure and full disclosure estimates are small). In column 3 there is limited evidence of gains in either fully assuming homosexuality (9.2%) or fully hiding it (8.2%). Therefore, there is limited evidence indicating possible taste for discrimination in the samples considered.

8 Conclusion

In this research I provide the first estimation of wage gaps, for Portugal, on primary employment by virtue of sexual orientation. The most remarkable result is the evidence of no statistically significant wage gaps. This finding is based on the most trustable method employed in the analysis: exact matching. Regression methods were also employed using a random and repre-

sentative sample of heterosexuals as counterfactuals. From the Blinder-Oaxaca decomposition, gay characteristics attenuate the negative wage gap whereas, for lesbians, the observed covariates induce a more severe wage gap. There is also evidence (though imprecise) that taste for discrimination may exist in individuals disclosing their homosexuality. Nonetheless, this indication of discrimination is suggested only by small differences among the regression coefficient estimates for disclosure. In addition, samples disparities between full and partial disclosure estimates raise further questions on whether there is in fact discrimination.

Although the previously described empirical results point towards a no discrimination conclusion, the tiny evidence of discrimination, together with the non-negligible proportion of homosexuals in fear of disclosing their sexual orientation, demand further investigation in the topic. Moreover, because the empirical results are only a particular conclusion for the samples used in this study (individuals more comfortable in disclosing their sexual orientation in Inquérito ao Emprego or more affirmative and active individuals from the web-survey that are engaged in social networks or connected to LGBTI institutions), a generalisation of the no discrimination conclusion requires further analysis of other samples.

Howbeit, even though this study provides some evidence of no discrimination, Portugal still has a long way to go on integrating homosexuals in the labour market, as suggested with the proportion of individuals that claim to suffer discrimination. Accordingly, this research is only a starting point. Confident evaluations on the full integration of homosexual individuals in the Portuguese labour market require further extensive research covering the various angles of homosexuality discrimination. Tackling the existing self-reported discrimination, together with studying employment (widely covered in foreign literature) and wage gaps including bonuses and promotions are important following steps from here, which I hope to have stimulated by kicking-of the topic.

9 References

Ahmed, Ali M., and Mats Hammarstedt. 2010. "Sexual orientation and earnings: A register data-based approach to identify homosexuals." *Journal of Population Economics* 23(3): 835-849.

Ahmed, Ali, Lina Andersson, and Mats Hammarstedt. 2013. "Sexual orientation and full-time monthly earnings, by public and private sector: evidence from Swedish register data." *Review of Economics of the Household* 11(1): 83-108.

Allegretto, Sylvia A., and Michelle M. Arthur. 2001. "An empirical analysis of homosexual/heterosexual male earnings differentials: unmarried and unequal?." *ILR Review* 54(3): 631-646.

Antecol, Heather, Anneke Jong, and Michael Steinberger. 2008. "The sexual orientation wage gap: The role of occupational sorting and human capital." *ILR Review* 61(4): 518-543.

Arrow, Kenneth. 1973. "The theory of discrimination." In *Discrimination in labor markets*, ed. Orley Ashenfelter and Albert Rees, 3-33. Princeton, NJ: Princeton University Press.

Badgett, M. L. 1995. The wage effects of sexual orientation discrimination.*ILR Review* 48(4): 726-739.

Badgett, MV Lee. 1996. "Employment and sexual orientation: Disclosure and discrimination in the workplace." *Journal of Gay & Lesbian Social Services* 4(4): 29-52.

Baert, S. (2017). Hiring a Gay Man, Taking a Risk?: A Lab Experiment on Employment Discrimination and Risk Aversion. *Journal of homosexuality*, 1-17.

Berg, Nathan, and Donald Lien. 2002. "Measuring the effect of sexual orientation on income: Evidence of discrimination?." *Contemporary economic policy* 20(4): 394-414.

Becker, Gary S. 1957. *The economics of discrimination*. University of Chicago press.

Becker, Gary S. 1976. *The economic approach to human behavior*. University of Chicago press.

Becker, Gary S. 1981. *A Treatise on the Family*. Harvard university press.

Black, Dan, Gary Gates, Seth Sanders, and Lowell Taylor. 2000. "Demographics of the gay and lesbian population in the United States: Evidence from available systematic data sources." *Demography* 37(2): 139-154.

Black, Dan A., Hoda R. Makar, Seth G. Sanders, and Lowell J. Taylor. 2003. "The earnings effects of sexual orientation." *ILR Review* 56(3): 449-469.

Black, Dan, A., Seth G. Sanders, and Lowell J. Taylor. 2007. "The Economics of Lesbian and Gay Families." *Journal of Economic Perspectives* 21(2): 53-70.

Blackwell, Matthew, Stefano M. Iacus, Gary King, and Giuseppe Porro. 2009. "cem: Coarsened exact matching in Stata." *The Stata Journal 9* 4: 524-546.

Blandford, John M. 2003. "The nexus of sexual orientation and gender in the determination of earnings." *ILR Review* 56(4): 622-642.

Buser, Thomas, Lydia Geijtenbeek, and Erik Plug. 2015. "Do gays shy away from competition? Do lesbians compete too much?." No 9382, IZA Discussion Papers, Institute for the Study of Labor (IZA).

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: methods and applications*. Cambridge university press.

Carpenter, Christopher. 2004. "New evidence on gay and lesbian household incomes." *Contemporary economic policy* 22(1): 78-94.

Carpenter, Christopher S. 2005. "Self-reported sexual orientation and earnings: Evidence from California." *ILR Review* 58(2): 258-273.

Carpenter, Christopher S. 2007. "Revisiting the income penalty for behaviorally gay men: Evidence from NHANES III." *Labour economics* 14(1): 25-34.

Carpenter, Christopher. 2008. "Sexual orientation, income, and non-pecuniary economic outcomes: New evidence from young lesbians in Australia." *Review of Economics of the Household* 6(4): 391-408.

Dalia Research. 2016. "Counting the LGBT Population: 6% of Europeans Identify as LGBT." Accessed December 18. https://daliaresearch.com/counting-the-lgbt-population-6-of-europeans-identify-as-lgbt/

Drydakis, Nick. 2009. "Sexual orientation discrimination in the labour market." *Labour Economics* 16(4): 364-372.

ILGA Europe. 2017. "Rainbow Europe: Portugal." Accessed December 18. https://rainbow-europe.org

Jann, Ben. 2008. "The Blinder-Oaxaca decomposition for linear regression models." *The Stata Journal* 8(4): 453-479.

Jepsen, Christopher, and Lisa K. Jepsen. 2015. "LaborMarket Specialization within SameSex and Difference-Sex Couples." *Industrial Relations: A Journal of Economy and Society* 54(1): 109-130.

Klawitter, Marieka M., and Victor Flatt. 1998. "The effects of state and local antidiscrimination policies on earnings for gays and lesbians." *Journal of policy analysis and management* 658-686.

La Nauze, Andrea. 2015. "Sexual orientationbased wage gaps in Australia: The potential role of discrimination and personality." *The Economic and Labour Relations Review* 26(1): 60-81.

Lemieux, Thomas. 2006. "The Mincer equation thirty years after schooling, experience, and earnings." In *Jacob Mincer A Pioneer of Modern Labor Economics*, ed. Shoshana Grossbard, 127-145. New York: Springer.

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings*. National Bureau of Economic Research, Inc.

Mincer, Jacob. 1975. "Education, experience, and the distribution of earnings and employment: an overview." In *Education, income, and human behavior*, ed. F. Phelps, Edmund S. 1972. "The statistical theory of racism and sexism." *The American Economic Review* 62(4): 659-661.

Plug, Erik, and Peter Berkhout. 2004. "Effects of sexual preferences on earnings in the Netherlands." *Journal of Population Economics* 17(1): 117-131.

Plug, Erik, Dinand Webbink, and Nick Martin. 2014. "Sexual orientation, prejudice, and segregation." *Journal*

of Labor Economics 32(1): 123-159.

Ragins, Belle Rose, and John M. Cornwell. 2001. "Pink triangles: antecedents and consequences of perceived workplace discrimination against gay and lesbian employees." *Journal of applied psychology* 86(6): 1244.

Sabia, Joseph J. 2014. "Sexual Orientation and Wages in Young Adulthood: New Evidence from Add Health." *ILR Review* 67(1): 239-267.

Sabia, Joseph J., Mark Wooden, and Thanh Tam Nguyen. 2017. "Sexual Identity, SameSex Relationships, and Labour Market Dynamics: New Evidence from Longitudinal Data in Australia." *Southern Economic Journal* 83(4): 903-931.

Thomas Juster, 71-94. Cambridge, Massachusetts: National Bureau of Economic Research.

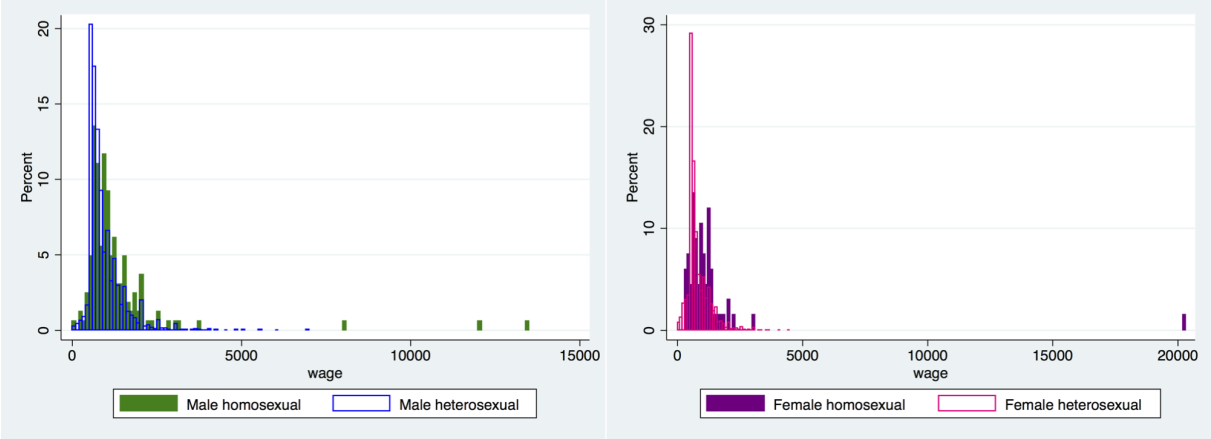
Tilcsik, Andrs. 2011. "Pride and prejudice: Employment discrimination against openly gay men in the United States." *American Journal of Sociology* 117(2): 586-626.

Valfort, Marie-Anne. 2017. "LGBTI in OECD Countries: A Review". No 198, OECD Working Paper.

Weichselbaumer, Doris. 2015. "Testing for discrimination against lesbians of different marital status: A field experiment." Industrial Relations: A Journal of Economy and Society 54(1): 131-161.

10 Appendices

Figure 2: Histogram of homosexual and heterosexual wages distribution with outliers

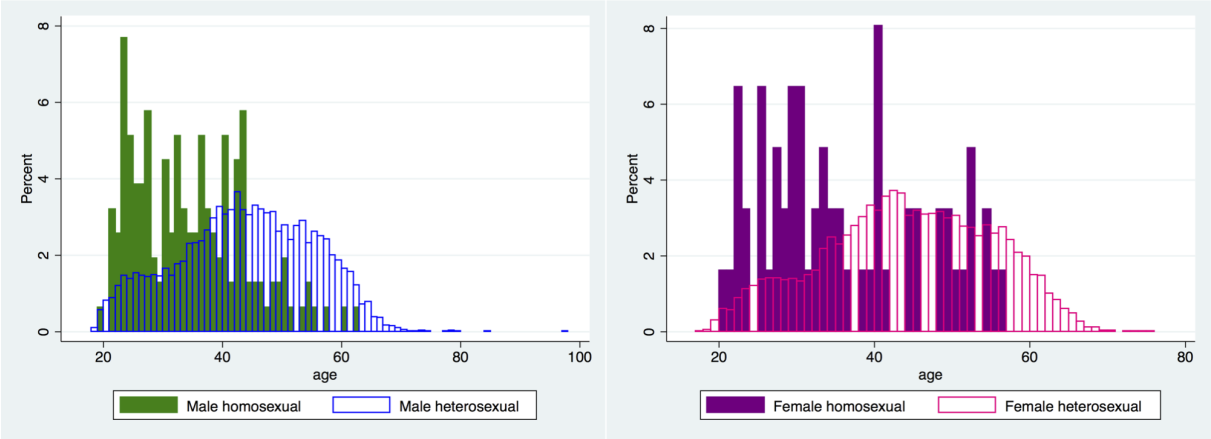


Source: Homosexual observations form the web-survey and heterosexual observations from Inquérito ao Emprego Q2 2017.

Figure 3: Histogram of homosexual and heterosexual years of education without outliers

Source: Homosexual observations form the web-survey and heterosexual observations from Inquérito ao Emprego Q2 2017.

Figure 4: Histogram of homosexual and heterosexual age distribution without outliers



Source: Homosexual observations form the web-survey and heterosexual observations from Inquérito ao Emprego Q2 2017.

Table 7: Percentage of observations per work region

Work Region/Sample	Inquérito ao Emprego Males		Web-Survey Males		Inquérito ao Females		Web-Survey Females	
	Straights	Gays	Straights	Gays	Straights	Lesbians	Straights	Lesbians
North	26.5%	13.9%	8.1%	12.8%	24.9%	-	7.5%	9.7%
Center	15.8%	18.8%	5.4%	4.5%	15.5%	44.9%	8.6%	9.7%
Lisbon	18.8%	55.5%	81.1%	66.7%	19.9%	32.7%	79.6%	67.7%
Alentejo	11.3%	2.0%	-	1.3%	11.3%		1.1%	1.6%
Algarve	9.7%	1.0%	5.4%	10.3%	11.1%	22.5%	1.6%	8.1%
Azores	9.4%	-	-	3.9%	8.4%	-	1.6%	1.6%
Madeira	8.5%	8.9%	-	0.6%	9.0%	-	-	1.6%

Source: Homosexual observations form the web-survey and heterosexual observations from Inquérito ao Emprego Q2 2017.

Table 8: Percentage of observations per work industry

Activity/Sample	Inquérito ao Emprego Males		Web-Survey Males		Inquérito ao Emprego Females		Web-Survey Females	
	Straights	Gays	Straights	Gays	Straights	Lesbians	Straights	Lesbians
A. Agriculture, hunting, forestry and fishing	3.6%	-	4.1%	0.6%	1.5%	-	1.6%	-
B. Extractive Industries	0.8%	1.0%	-	-	0.1%	-	-	-
C. Manufacturing	19.5%	3.0%	-	5.0%	13.5%	6.1%	2.7%	1.6%
D. Electricity, gas, hot and cold water and air conditioning	1.0%	-	2.7%	0.6%	0.2%	-	-	-
E. Water collection, treatment and distribution; sanitation and waste management	1.5%	1.0%	-	-	0.3%	-	0.5%	-
F. Construction	11.9%	1.0%	1.4%	0.6%	0.8%	-	1.1%	1.6%
G. Wholesale and retail trade; repair of vehicles	13.8%	25.7%	2.7%	3.9%	13.1%	32.7%	4.3%	4.8%
H. Transportation and storage	7.1%	9.9%	1.4%	0.6%	1.6%	2.0%	-	1.6%
I. Accommodation, catering and similar	5.1%	22.8%	2.7%	8.3%	8.5%	20.4%	1.6%	8.1%
J. Information and communication activities	2.5%	4.0%	4.1%	8.3%	1.3%	-	3.8%	6.5%
K. Financial and insurance activities	2.5%	9.9%	37.8%	5.8%	2.0%	4.1%	17.2%	4.8%
L. Real estate activities	0.4%	1.0%	-	2.6%	0.6%	-	1.6%	-
M. Consulting, scientific, technical and similar activities	2.2%	10.9%	16.2%	13.5%	3.4%	-	11.3%	19.4%
N. Administrative and support services activities	3.8%	1.0%	2.7%	10.9%	3.0%	12.2%	6.5%	8.1%
O. Public administration and defence; compulsory social security	13.0%	-	6.8%	3.2%	7.8%	20.4%	3.2%	6.5%
P. Education	5.1%	1.0%	9.5%	11.5%	17.6%	-	21.0%	9.7%
Q. Human health and social support activities	3.6%	4.0%	-	13.5%	18.4%	2.0%	12.4%	17.7%
R. Arts, entertainment, sports and recreational activities	1.4%	-	-	3.2%	1.2%	-	2.7%	6.5%
S. Other service activities	1.1%	4.0%	8.1%	7.7%	2.0%	-	7.5%	3.2%
T. Activities of households employing domestic staff and household production activities for own use	0.1%	-	-	-	3.4%	-	1.1%	-
U. Activities of international organizations and other extra-territorial institutions	0.1%	-	-	-	0.1%	-	-	-

Source: Homosexual observations form the web-survey and heterosexual observations from Inquérito ao Emprego Q2 2017.

Table 9: Percentage of observations per occupation

Occupation/Sample	Inquérito ao Emprego Males		Web-Survey Males		Inquérito ao Emprego Females		Web-Survey Females	
	Straights	Gays	Straights	Gays	Straights	Lesbians	Straights	Lesbians
0. Military	1.2%	-	-	-	0.1%	-	-	-
1. Representatives of the legislative and executive bodies, officers, directors and executive managers	4.1%	-	2.7%	5.8%	2.3%	20.4%	3.2%	3.3%
2. Specialists in intellectual and scientific activities	12.7%	27.7%	46.0%	42.2%	21.8%	8.2%	37.6%	41.0%
3. Intermediate level technicians and professions	13.3%	14.9%	35.8%	20.1%	12.0%	14.3%	31.2%	31.2%
4. Administrative staff	7.0%	4.0%	8.6%	10.4%	11.8%	10.2%	19.4%	11.5%
5. Personal services, security and safety, and vendors	13.4%	40.6%	2.7%	11.0%	23.7%	40.8%	5.9%	8.2%
6. Farmers and skilled workers in agriculture, fishery and forestry	3.4%	-	1.4%	-	0.7%	-	-	-
7. Skilled workers of industries, construction and craftsmen	22.1%	4.0%	1.4%	2.0%	4.3%	2.0%	0.5%	1.6%
8. Operators of plant and machinery; assembly workers	14.1%	4.0%	1.4%	0.7%	5.8%	2.0%	-	-
9. Unskilled workers	8.8%	5.0%	-	7.8%	17.6%	2.0%	2.2%	3.3%

Source: Homosexual observations form the web-survey and heterosexual observations from Inquérito ao Emprego Q2 2017.

Table 10: OLS estimation impact of homosexuality on wages for partnered individuals

Variables/Samples	Homosexual & Straigth (IE)	
	(1)	(2)
Males		
Gay	0.105	**0.043**
	(0.11)	(0.06)
R-squared	0.002	0.616
Controls		YES
Observations	100163	87338
Females		
Lesbian	0.103	**0.043**
	(0.13)	(0.08)
R-squared	0.005	0.750
Controls		YES
Observations	99990	91297

Source: Columns 1 and 2 use partnered homosexual and heterosexual observations from Inquérito ao Emprego, from Q1 2011 to Q2 2017. The controls used are: nationality, marital status, level of education, experience, work region, full time status, industry, occupation, hours worked, job search, size of the workplace, work contract, professional situation and year-quarter fixed effects. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 11: OLS estimation impact on wages of having a same-sex partner in the household

Variables/Samples	(1) Gay (WS) & Straight Male (IE)	(2) Gay & Straight Male (WS)	(3) Lesbian (WS) & Straight Female (IE)	(4) Lesbian & Straight Female (WS)
Gay	0.056	0.064	-	-
	(0.04)	(0.06)		
Lesbian	-	-	0.043	0.060
			(0.05)	(0.08)
Same-Sex Cohabitor	**-0.036**	**-0.191***	**0.013**	**-0.023**
	(0.07)	(0.11)	(0.11)	(0.13)
R-squared	0.602	0.704	0.724	0.718
Controls	YES	YES	YES	YES
Observations	4940	205	5595	222

Source: Columns 1 and 2 use homosexual observations from the web-survey and heterosexual observations from Inquérito ao Emprego Q2 2017. Columns 3 and 4 homosexual and heterosexual observations form the web-survey. Each model uses either only male observations or female observations. The controls used are: nationality, marital status, level of education, experience, work region, full time status, industry, occupation, hours worked, job search, size of the workplace, work contract and professional situation. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 12: Blinder-Oaxaca Decomposition Results for partnered individuals

	Straight Male vs Gay	Straight Female vs Lesbian
Variables/Samples	(1) Gay & Straight (IE)	(2) Lesbian & Straight (IE)
Total Log Monthly Wage Gap	0.110	0.114
	(0.11)	(0.11)
Differences in characteristics	0.078	-0.28**
	(0.12)	(0.11)
Differences in coefficients	**0.018**	**0.104**
	(0.07)	(0.11)
Interaction of both	0.013	0.291
	(0.11)	(0.11)
Controls	YES	YES
Straight Observations	100062	99945
Homosexual Observations	101	45

Source: The dependent variable is the logarithm of monthly wages. Columns 1 and 2 use homosexual and heterosexual observations living with their partner in the same household from Inquérito ao Emprego, from Q1 2011 to Q2 2017. The controls used are: level of education, experience, full time status, and year and quarter fixed effects. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$